Chinese Journal of Aeronautics, (2014),27(1): 110–116

Chinese Society of Aeronautics and Astronautics
& Beihang University
Chinese Journal of Aeronautics
cja@buaa.edu.cn
www.sciencedirect.com

Weighted Marginal Fisher Analysis with Spatially Smooth for aircraft recognition

Wei Zhenzhong *, Liu Chang, Li Nan

Precision Opto-mechatrionics Technology, Key Laboratory of Education Ministry, Beihang University, Beijing 100191, China

Received 26 November 2012; revised 28 May 2013; accepted 24 September 2013
Available online 19 December 2013

KEYWORDS

Aircraft dataset;
Aircraft recognition;
Graph Embedding;
Invariant feature;
Laplacian operator;
Subspace learning

Abstract Due to limitations to extract invariant features for recognition when the aircraft presents various poses and lacks enough samples for training, a novel algorithm called Weighted Marginal Fisher Analysis with Spatially Smooth (WMFA-SS) for extracting invariant features in aircraft recognition is proposed. According to the Graph Embedding (GE) framework, Heat Kernel function is firstly introduced to characterize the interclass separability when choosing the weights of penalty graph. Furthermore, Laplacian penalty is applied to constraining the coefficients to be spatially smooth in this algorithm. Laplacian penalty is able to incorporate the prior information that neighboring pixels are correlated. Besides, using a Laplacian penalty can also avoid the singularity of Laplacian matrix of intrinsic graph. Once compact representations of the images are obtained, it can be considered as invariant features and then be performed in classification to recognize different patterns of aircraft. Real aircraft recognition experiments show the superiority of our proposed WMFA-SS in comparison to other GE algorithms and the current aircraft recognition algorithm; the accuracy rate of our proposed method is 90.00% for dataset BH-AIR1.0 and 99.25% for dataset BH-AIR2.0.

1. Introduction

Researches on recognizing 3D objects from 2D images are becoming increasingly popular.[1] Analyzing features of aircraft as typical 3D dynamic objects and identifying their patterns are always attracted by many scholars. Concerning the issues for diversities of aircraft's poses such as variation of 3D scale, transformation and rotation, it makes recognition of aircraft images quite difficult.

The current methodologies of recognition for aircraft objects are to extract invariant features based on shape information and then to discriminate aircraft patterns combined with various classifiers. Invariant features, such as affine moment, wavelet moment, sift feature, etc. have been extensively adopted as effective feature extraction means, which have appeared pros and cons in certain specific applications. Among invariant features mentioned above, different features appear different tolerance for invariance on condition of images taken by various poses of aircraft. Flusser[2] proposed the affine moment, which is able to keep invariant to image distortion or twisting caused by small aircraft's roll angle and pitch angle variation; wavelet

* Corresponding author. Tel.: +86 10 82338768.
E-mail addresses: zhenzhongwei@buaa.edu.cn (Z. Wei), vincentliubuaa@gmail.com (C. Liu).

Peer review under responsibility of Editorial Committee of CJA.

ELSEVIER Production and hosting by Elsevier

1000-9361 © 2014 Production and hosting by Elsevier Ltd. on behalf of CSAA & BUAA. Open access under CC BY-NC-ND license.
http://dx.doi.org/10.1016/j.cja.2013.12.014

moment[3] is capable of keeping invariant to aircraft's scaling variation; SIFT feature[4] can keep invariant to affine variation and resists noise quite well. However, one kind of invariant features can only satisfy recognition requirement limited in a specific circumstance and might not acquire a satisfying results shifting into another circumstance. Therefore, it is not feasible to apply single kind of features combining classifiers to construct an aircraft image recognition system that acquires high recognition rate on condition of aircraft's pose diversity.

According to the principle of integration, it is feasible to construct a more general aircraft image recognition system that can effectively acquire higher recognition rate on condition of different circumstances if invariant features are combined based on different rules. Zhu[5] proposed a method of aircraft recognition based on Multiple Classifier Fusion with Multiple Invariants (MCF-MI), which fused four different aircraft image features: affine moment, Zernike moment, wavelet moment and gradient module of SIFT feature descriptor and combined support vector machine to construct four kinds of classifiers. Moreover, an adaptive weighted voting method is adopted to carry out multiple classifier fusion for improving aircraft recognition rate, which is higher than the recognition rates using the classifiers constructed with single invariant features. However, it cannot satisfy the real-time requirement in the process of aircraft recognition because the choice of invariant features and the fusion of invariant features and classifiers are determined by different aircraft image conditions. Specifically, under the conditions of wide angle rotation of aircrafts, invariant features suffer trivial-solutions, leading to a dramatic a reduction in recognition accuracy.

Recently, there are considerable interests into visual analysis and dimension reduction. One hopes that estimating geometrical and topological properties of the manifold from random points lies on this unknown sub-manifold. Along this direction, many researchers[6–8] proposed lots of subspace learning algorithms to explore the local geometric structure embedded in high dimensional data. Some popular ones include Locality Preserving Projection (LPP),[9] Neighborhood Preserving Embedding (NPE)[10] and Marginal Fisher Analysis (MFA).[11] When using these methods, we usually represent an image of size $m_1 \times m_2$ pixels by a vector in an $m_1 \times m_2$ dimensional space. Although various poses of aircraft lead to huge difference in the same pattern of aircraft images and the dimensionality of data is extremely high when aircraft images are represented as vectors, there might be an intrinsic properties of the high manifold embedded in a low sub-manifold, which leads us to consider methodologies of subspace learning to extract invariant features from high-dimensional aircraft manifolds.

In this paper, we introduce a method called Weighted Marginal Fisher Analysis with Spatially Smooth for extracting invariant features in aircraft recognition. Based on the Graph Embedding framework,[12] we firstly introduce Heat Kernel function[13] to characterize the relationship between different pattern of aircraft images. Furthermore, we use Laplacian penalty[14] to constrain the coefficients to be spatially smooth in this algorithm. Instead of considering the basis function as a $m_1 \times m_2$ dimensional vector, we consider it as a matrix, or a discrete function defined on a $m_1 \times m_2$ lattice. So the discretized Laplacian operator can be applied to measuring the smoothness of basis function along the horizontal and vertical directions. Because the discretized Laplacian operator is a finite difference approximation to the second derivative operator, we sum over all directions. The choice of Laplacian penalty allows us to incorporate the prior information that neighboring pixels are correlated. Besides, using a Laplacian penalty can also avoid the singularity of Laplacian matrix of intrinsic graph. After we acquire representations of images in the subspace, we can consider them as invariant features and then perform classification to recognize different patterns of aircraft.

The method presents two essential characteristics:

(1) Before acquiring the optimized projections that maximize the interclass separability represented by penalty graph and minimize the intraclass compactness represented by intrinsic graph, the choice of Heat Kernel function as weights of penalty graph can characterize the interclass separability better than simple-minded function used by MFA in neighborhood relationships.
(2) Although PCA as a pre-projection to avoid the singularity of Laplacian matrix of intrinsic graph works very well in various object recognitions like face recognition,[15] it failed in aircraft recognition for huge difference caused by pose variation between the same pattern of aircrafts. Therefore, we introduce Laplacian penalty as the regularization to avoid the singularity of Laplacian matrix of intrinsic graph and to avoid over-fitting[16] because the number-of-dimensions to the sample-size ratio is too high.

The rest of the paper is structured as follows. In Section 2, we define the feature extraction as a subspace learning representation in aircraft recognition and provide a brief review of the Graph Embedding framework and Laplacian smoothing. Section 3 introduces our proposed algorithm called Weighted Marginal Fisher Analysis with Spatially Smooth. The extensive experimental results are presented in Section 4. Finally, we provide some conclusive remarks and suggestions for future work in Section 5.

2. Problem definition and theoretical background

2.1. Problem definition

For a general classification problem, each image of the aircraft sample set is rearranged from size of $m_1 \times m_2$ to a vector $\boldsymbol{x}_i \in \mathbf{R}^m$, where $m = m_1 \times m_2$. The aircraft sample set for model training can be represented as a matrix $\boldsymbol{X} = [\boldsymbol{x}_1\ \boldsymbol{x}_2\ \dots\ \boldsymbol{x}_N]$, where N is the sample number and m the data dimension. For supervised learning problems, the class label of sample $\boldsymbol{x}_i$ is assumed to be $c_i = \{1, 2, \dots, N_c\}$, where N_c is the number of classes. We also let π_c and n_c denote the index set and number of samples belonging to the class, respectively. The linear feature extraction is to find a mapping matrix $\boldsymbol{A} = [\boldsymbol{a}_1\ \boldsymbol{a}_2\ \dots\ \boldsymbol{a}_l] \in \mathbf{R}^{m \times l}$, which transforms sample $\boldsymbol{x}_i$ to a specific representation $\boldsymbol{y}_i$, where, typically, l m and $\boldsymbol{y}_i = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{x}_i$.

2.2. A brief review of Graph Embedding framework

Let $G = \{\boldsymbol{X}, \boldsymbol{W}\}$ be an undirected weighted graph with vertex set $\boldsymbol{X}$ and similarity matrix $\boldsymbol{W} \in \mathbf{R}^{N \times N}$. Each element of the real symmetric matrix $\boldsymbol{W}$ measures, for a pair of vertices, its similarity, which might be negative. In this work, the Graph

Embedding of the graph G is defined as an algorithm to find the relationships of the desired low-dimensional vector representations between the vertices of G that best characterize the similarity relationship between the vertex pairs in G.

For simplify exposition, we take the one-dimensional case and let vector $\boldsymbol{y} = [y_1\ y_2 \ldots y_N]^{\mathrm{T}}$ be the representation of vertex $\boldsymbol{X} = [\boldsymbol{x}_1\ \boldsymbol{x}_2 \ldots \boldsymbol{x}_N]$. If the mapping is linear, $y_i = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}_i$. The objective function of Graph Embedding is given as follows:

$$\boldsymbol{a}^* = \arg \min_{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{XBX}^{\mathrm{T}}\boldsymbol{a}=d} \boldsymbol{a}^{\mathrm{T}}\boldsymbol{XLX}^{\mathrm{T}}\boldsymbol{a} \tag{1}$$

where d is a constant and $\boldsymbol{B}$ is the constraint matrix; $\boldsymbol{B}$ is typically a diagonal matrix for scale normalization and may also be the Laplacian matrix of a penalty graph G^{p}, and $\boldsymbol{L}$ is the Laplacian matrix of an intrinsic graph G. Laplacian matrix of the intrinsic graph and the penalty graph can be defined as follows:

$$\begin{cases} \boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W},\ D_{ii} = \sum_{i \neq j} W_{ij},\ \forall i \in \mathbf{N} \\ \boldsymbol{B} = \boldsymbol{D}^{\mathrm{p}} - \boldsymbol{W}^{\mathrm{p}},\ D^{\mathrm{p}}_{ii} = \sum_{i \neq j} W^{\mathrm{p}}_{ij},\ \forall i \in \mathbf{N} \end{cases} \tag{2}$$

where $\boldsymbol{D} = \mathrm{diag}(D_{11}, D_{22}, \cdots, D_{NN})$.

Given different similarity matrices $\boldsymbol{W}$, function Eq. (1) can be represented as different subspace learning algorithms (PCA, LDA, NPE and MFA, etc.).

2.3. Laplacian smoothing

Let f be a function defined on a region of interest $\Omega \in \mathbf{R}^d$. The Laplacian operator $\mathcal{L}$ is defined as[18]

$$\mathcal{L}f(\boldsymbol{t}) = \sum_{j=1}^{d} \frac{\partial^2 f}{\partial \boldsymbol{t}_j^2} \tag{3}$$

here, d defines the dimensionality of the signal.

The Laplacian penalty function, denoted by J, is defined by

$$J(f(\boldsymbol{t})) = \int_{\Omega} [\mathcal{L}f(\boldsymbol{t})]\mathrm{d}\boldsymbol{t} \tag{4}$$

Intuitively, $J(f(\boldsymbol{t}))$measures the smoothness of the function f over the region Ω. In this paper, our interest lies in image. An image is intrinsically a two-dimensional signal. Therefore, we take d to be 2 in the proposed algorithm.

3. Weighted Marginal Fisher Analysis with Spatially Smooth (WMFA-SS)

3.1. Constructing graph of intraclass and graph of interclass

If $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ represent two images of the same aircraft with two different poses, we design the intrinsic graph that characterizes the intraclass relationship based on Graph Embedding framework. In the intrinsic graph, a vertex pair is connected if one vertex $\boldsymbol{x}_i$ is among the k_1-nearest neighbors of the other $\boldsymbol{x}_j$ and the elements of the pair belong to the same class. We choose the same weights to describe the connected relationship to reduce the intraclass's difference. The similarity matrix of the intrinsic graph is given as

$$W_{i,j} = \begin{cases} 1, & \text{if } i \in N^{+}_{k_1}(j) \text{ or } j \in N^{+}_{k_1}(i) \\ 0, & \text{else} \end{cases} \tag{5}$$

here, $N^{+}_{k_1}(i)$ indicates the index set of the k_1-nearest neighbors of the sample $\boldsymbol{x}_i$ in the same class.

If $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ represent two images of the two classes of aircraft with possible similar poses, we design the penalty graph that characterizes the interclass relationship based on Graph Embedding framework. In the penalty graph, for each class, the k_2-nearest vertex pairs in which one element is in-class and the other is out-of-class are connected. We choose Heat Kernel function to describe the connected relationship to increase the interclass's divergence. The similarity matrix of the penalty graph is defined as

$$W^{\mathrm{p}}_{i,j} = \begin{cases} \mathrm{e}^{-\|\mathbf{x}_i - \mathbf{x}_j\|^2/\sigma}, & \text{if } (i,j) \in P_{k_2}(c_i) \ or \ (i,j) \in P_{k_2}(c_i) \\ 0, & \text{else} \end{cases} \tag{6}$$

here, $P_{k_2}(c_i)$ is a set of data pairs that are the k_2-nearest pairs among the set $\{(i,j), i \in \pi_c, j \notin \pi_c\}$, and σ is a constant.

By following the Graph Embedding formulation Eq. (1), the object function of WMFA-SS is characterized as

$$\begin{aligned} \boldsymbol{a}^* &= \arg \min_{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{XBX}^{\mathrm{T}}\boldsymbol{a}=d} \boldsymbol{a}^{\mathrm{T}}\boldsymbol{XLX}^{\mathrm{T}}\boldsymbol{a} = \arg\max_{\boldsymbol{a}} \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{XBX}^{\mathrm{T}}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{XLX}^{\mathrm{T}}\boldsymbol{a}} \\ &= \arg\max_{\boldsymbol{a}} \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{D}^{\mathrm{p}} - \boldsymbol{W}^{\mathrm{p}})\boldsymbol{X}^{\mathrm{T}}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{D} - \boldsymbol{W})\boldsymbol{X}^{\mathrm{T}}\boldsymbol{a}} \\ &= \arg\max_{\boldsymbol{a}} \frac{\sum_i \sum_{(i,j)\in P_{k_2}(c_i) \text{ or } (i,j)\in P_{k_2}(c_i)} W^{\mathrm{p}}_{i,j} \|\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}_i - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}_j\|^2}{\sum_i \sum_{i\in N^{+}_{k_1}(j) \text{ or } j \in N^{+}_{k_1}(i)} \|\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}_i - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}_j\|^2} \end{aligned} \tag{7}$$

here, $W^{\mathrm{p}}_{i,j}$ is the value of similarity matrix as defined in Eq. (6). According to the formation in Eq. (5), we could neglect $W^{\mathrm{p}}_{i,j}$ because $W_{i,j} = 1$ if $\boldsymbol{x}_i$ is among the k_1-nearest neighbors of the other $\boldsymbol{x}_j$ or $\boldsymbol{x}_j$ is among the k_1-nearest neighbors of the other $\boldsymbol{x}_i$.

3.2. Objective function of WMFA-SS

As we described previously, $m_1 \times m_2$ aircraft images can be represented as vectors in $\mathbf{R}^m$, $m = m_1 \times m_2$. Let $\boldsymbol{a}_i \in \mathbf{R}^m$ be the basis vectors obtained by WMFA-SS. Without loss of generality, $\boldsymbol{a}_i$ can also be considered as functions defined on a $m_1 \times m_2$ lattice.

For an aircraft image, the region of interest Ω is a two-dimensional rectangle, for notational convenience of Ω, we take it to be 2. A lattice is defined on Ω as follows. Let $\boldsymbol{h} = [h_1\ h_2]$ where $h_1 = 1/m_1$ and $h_2 = 1/m_2$. $\Omega_{\boldsymbol{h}}$ consists of the set of two-dimensional vectors $\boldsymbol{t}_i = [t_{i1}\ t_{i2}]$ with $t_{ij} = (i_j - 0.5)\ h_j$ for $1 \leqslant i_j \leqslant m_j$ and $1 \leqslant j \leqslant 2$. There are a total of $m = m_1 \times m_2$ grid points in this lattice. Let $\boldsymbol{D}_j$ be an $m_1 \times m_2$ matrix that yields a discrete approximation to $\partial^2/\partial \boldsymbol{t}_j^2$. Thus if $\boldsymbol{u} = [u(t_1)\ u(t_2) \ldots u(t_{m_j})]$ is an m_j -dimensional vector which is a discretized version of a function $u(\boldsymbol{t})$, then $\boldsymbol{D}_j$ has the property that

$$\boldsymbol{D}_j \boldsymbol{u}(t_{ij}) \approx \frac{\partial^2 u(t_{ij})}{\partial \boldsymbol{t}_j^2} \tag{8}$$

For $i = 1, 2, \ldots, m_j$. There are many possible choices of $\boldsymbol{D}_j$.[17] In this work, we apply the modified Neuman discretization[18]

$$D_j = \frac{1}{h_j^2}\begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{bmatrix} \tag{9}$$

Given $\boldsymbol{D}_j$, a discrete approximation $\boldsymbol{\Delta}$ for two-dimensional Laplacian $\mathcal{L}$ is the $m \times m$ matrix

$$\boldsymbol{\Delta} = \boldsymbol{D}_1 \otimes \boldsymbol{I}_2 + \boldsymbol{I}_1 \otimes \boldsymbol{D}_2 \tag{10}$$

where $\boldsymbol{I}_j$ is $m_j \times m_j$ identity matrix for $j = 1, 2$. $\otimes$ is the kronecker product.[19]

For an $m_1 \times m_2$ dimensional vector $\boldsymbol{a}$, it is easy to check that $||\boldsymbol{\Delta} \cdot \boldsymbol{a}||^2$ is proportional to the sum of the squared differences between the nearby grid points of $\boldsymbol{a}$ with its matrix form. It provides a measure of smoothness of $\boldsymbol{a}$ on the $m_1 \times m_2$ lattice.

Given the WMFA-SS structure with penalty Laplacian matrix $\boldsymbol{B}$ and intrinsic Laplacian matrix $\boldsymbol{L}$, the WMFA-SS is defined as

$$\arg\max_{\boldsymbol{a}} \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{a} + \alpha J(\boldsymbol{a})} \tag{11}$$

where the parameter $0 \leqslant \alpha \leqslant 1$ controls the smoothness of the estimator; when $\alpha = 0$, the WMFA-SS will reduce to the WMFA which totally ignores the spatial relationship between pixels of an image; when $\alpha \to \infty$, the algorithm will choose $\boldsymbol{a}$ spatially smoothest basis vector and totally ignore the manifold structure of the aircraft data. J is the discretized Laplacian regularization functional

$$J(\boldsymbol{a}) = ||\boldsymbol{\Delta} \cdot \boldsymbol{a}||^2 = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}\boldsymbol{a} \tag{12}$$

The vectors $\boldsymbol{a}_i$ ($i = 1, 2, \ldots, l$) that maximize the objective function Eq. (11) are given by the maximum eigenvalue solutions to the following generalized eigenvalue problem

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{a} = \lambda\left(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{a} + \alpha J(\boldsymbol{a})\right) \tag{13}$$

The algorithmic procedure of WMFA-SS is formally stated as follows:

(1) **Constructing the graph of intraclass compactness and graph of interclass separability.** In the intraclass compactness, using simple-minded function Eq. (5), that is, for each sample $\boldsymbol{x_i}$, set the similarity matrix $W_{ij} = W_{ji} = 1$ if $\boldsymbol{x_i}$ is among the k_1-nearest neighbors of $\boldsymbol{x_j}$ in the same class. In the interclass separability graph, using Heat Kernel function Eq. (6), for each class c, set the similarity matrix $W_{i,j}^{\mathrm{P}} = \mathrm{e}^{-||x_i - x_j||^2/\sigma}$ if the pair (i,j) is among the k_2-shortest pairs among the set $\{(i,j), i \in \pi_c, j \notin \pi_c\}$.

(2) **Computing the discrete approximation $\boldsymbol{\Delta}$ for two-dimensional Laplacian $\mathcal{L}$.** From the formulation Eqs. (9) and (10), the discrete approximation $\boldsymbol{\Delta}$ is
$\boldsymbol{\Delta} = \boldsymbol{D}_1 \otimes \boldsymbol{I}_2 + \boldsymbol{I}_1 \otimes \boldsymbol{D}_2$

(3) **Constructing the WMFA-SS object function.** Given the WMFA structure with penalty Laplacian matrix $\boldsymbol{B}$ and intrinsic Laplacian matrix $\boldsymbol{L}$, the WMFA-SS is defined as follows:

$$\arg\max_{\boldsymbol{a}} \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{a} + \alpha J(\boldsymbol{a})}$$

(4) **Embedding to l dimensional subspace.** Let basis vectors $\boldsymbol{a_i}$ ($i = 1, 2, \ldots, l$) that maximize the objective function Eq. (11) are given by the maximum eigenvalue solutions to the generalized eigenvalue problem Eq. (13). Let $\boldsymbol{A} = [\boldsymbol{a}_1\ \boldsymbol{a}_2\ \ldots\ \boldsymbol{a}_l]$ which is a $m \times l$ projection matrix. The samples can be embedded into l dimensional subspace by

$$\boldsymbol{X} \to \boldsymbol{Y} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}$$

The compact representations $\boldsymbol{Y} = [\boldsymbol{y}_1\ \boldsymbol{y}_2\ \ldots\ \boldsymbol{y}_N]$ embedded in this subspace are invariant features, whose number is N and dimension is l.

Finally, we represent the testing samples into the subspace learned from the training sample and then use k-NN classification to classify different patterns of aircraft in the l dimensional subspace.

4. Experimental results

4.1. Aircraft dataset and experiment settings

Due to the lack of 3D aircraft object data sample, we utilize 3ds Max software to build two aircraft image datasets – (BH-AIR-1.0 and BH-AIR-2.0) to evaluate the performance of our proposed algorithm.

The purpose of the first dataset is to demonstrate that the algorithm is still recognizable when the aircraft has big space rotation while the MCF-MI fails and at the same time verify the superiority of this algorithm by comparing with current GE algorithms. In this dataset, we choose four aircraft: F22, F111, Su27 and Mirage2000. Each aircraft includes 90 viewpoints, which defines 18 rotation viewpoints from −180° to 180° with every 20° interval and 5 viewpoints moving the aircraft 5 pixels in the up, down, left and right directions in the image (Fig. 1). We preprocess all these images with binary images and the same size 48×48 in dataset before we verify the experiment. The purpose of the second one is to verify the superiority of our algorithm by comparing with the current algorithm MCF-MI. In this dataset, we sample three aircraft: F22, F111, SU27 at different roll angles or pitch angles. Each

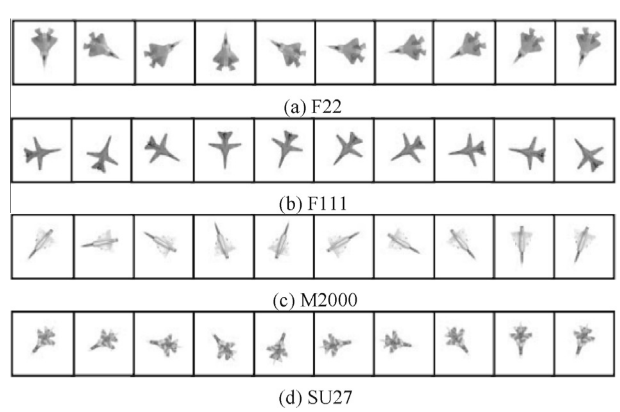


Fig. 1 Sample images from the aircraft database BH-AIR-1.0.

aircraft includes 81 viewpoints, which defines 9 viewpoints from $-20°$ to $20°$ with every $5°$ interval in roll angle and 9 viewpoints from $-20°$ to $20°$ with every $5°$ interval in pitch angle. We further scale these images with 6 different transformations with the ratio of 1/2, 3/4, 1, 5/4, 3/2, 7/4, 2 and finally acquire 567 images for each aircraft (Fig. 2). Thus, we acquire 1701 images in the second dataset and preprocess it the same as the first one. We summarize all of databases used in the experiments, which is listed in Table 1.

4.2. WMFA-SS versus GE algorithms

We first evaluate the performance of WMFA-SS in comparison to other classic GE algorithms: MFA, LDA on BH-AIR-1.0. For a fair comparison with these classic algorithms, we introduce Laplacian penalty as regularization into MFA and LDA for spatially smooth, which we called MFA-SS and LDA-SS. Moreover, we also transform data into PCA subspace first and then utilize WMFA, MFA and LDA, which we called PCA + WMFA, PCA + MFA, PCA + LDA.

Since results in subspace learning algorithms are affected by the number of training samples, dataset BH-AIR-1.0 was randomly split into 50 training samples and 40 test samples. As mentioned in Graph Embedding framework,[11] how to set parameters is still an open problem. We empirically set the parameters in WMFA-SS. Specifically, we utilize the cross validation to set the optimized parameters of intrinsic graph and penalty graph – k_1 is 10 and k_2 is 400. In penalty graph, the parameter σ in weight function is $(1/2N_{k_2})\sum||\boldsymbol{x}_i - \boldsymbol{x}_j||^2$, where $(|\boldsymbol{x}_i - |\boldsymbol{x}_j|)$ represents the nearest pairs among the k_2-nearest among the set $\{(i,j),\ i \in \pi_c,\ j \notin \pi_c\}$. The parameter of α is set as 0.1. A crucial problem for most of subspace learning algorithms is dimensionality estimation and often needs to be heuristically determined. Therefore, the performance usually varies with the number of dimensions. We show the best results of each algorithm obtained by 30 random splits of 50 training samples and 40 test samples in Fig. 3. As we can see, every algorithm reaching each peak recognition accuracy varies with the number of feature-dimensions. Finally, we report the experimental results of these six algorithms, which are listed in Table 2. The mean and standard deviations of recognition accuracy in 30 random splits are reported in the second column, and the numbers in parentheses are the corresponding feature dimension with the best recognition accuracy. The third column of this Table describes the best results in 30 random splits corresponding to certain feature dimensions as shown in parentheses.

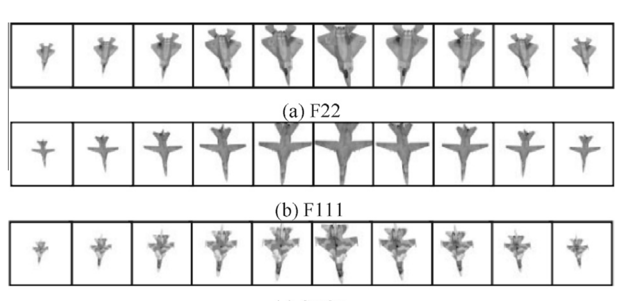


Fig. 2 Sample images from the aircraft database BH-AIR-2.0.

Table 1 Summary of the database (48 × 48) used in experiments.

	BH-AIR-1.0	BH-AIR-2.0
Number of image	360	1701
Number of aircraft	4	3
Number of image per aircraft	90	567

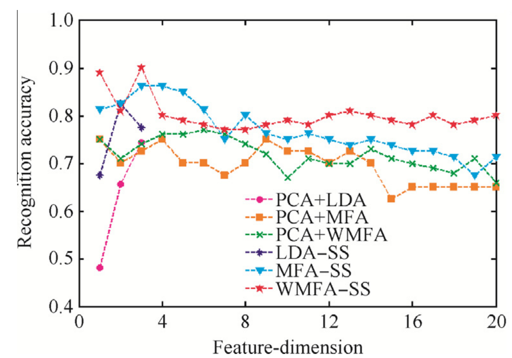


Fig. 3 Recognition accuracy versus feature-dimension on BH-AIR-1.0.

Table 2 Recognition accuracies of WMFA-SS versus GE algorithms.

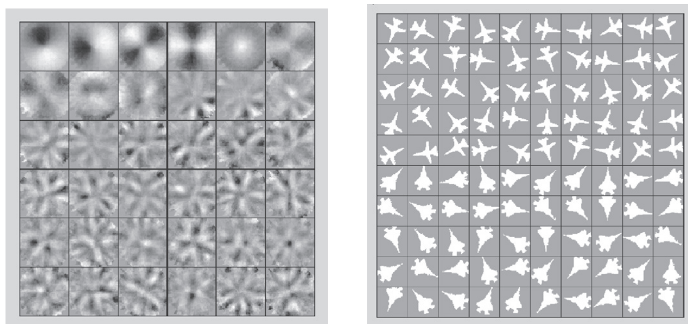
Method	Mean ± Std-dev (dimension)	Max (dimension)
PCA + LDA	70.1 ± 7.37 (3)	74.4 (3)
PCA + MFA	72.5 ± 0.57 (4)	75.0 (4)
PCA + WMFA	73.1 ± 1.34 (6)	77.0 (6)
LDA-SS	70.5 ± 0.49 (2)	82.5 (2)
MFA-SS	81.1 ± 0.58 (3)	86.3 (3)
WMFA-SS	85.4 ± 0.78 (3)	90.0 (3)

The main observations from the performance comparison include: WMFA-SS approach significantly outperforms other GE algorithms. The reason lies in the fact that the Heat Kernel function is introduced in the penalty graph describing better relationships between different kinds of aircraft; moreover, Laplacian penalty, as a regularization, takes into account the spatial relationship of an image, solves the problem of intrinsic similarity matrix's singularity and avoids over-fitting due to limited data. We could also observe that introducing a Laplacian penalty for spatial smooth in other GE algorithms works better in aircraft recognition than PCA + GE algorithms, which previously work very well in face recognitions.[11]

According to PCA + GE algorithms, we projected the original aircraft samples into PCA subspace, as shown in Fig. 4. The projection vectors cannot represent the aircraft shape because of different poses while can represent the mostly face profile, thus the accuracy of PCA + GE algorithms performs not as well as our proposed algorithm.

4.3. WMFA-SS versus MCF-MI

To evaluate the superiority of our algorithm by comparing it with the current aircraft recognition method MCF-MI, we



(a) Original database (b) The first 36 PCA eigenvectors

Fig. 4 Original database and the first 36 PCA eigenvectors.

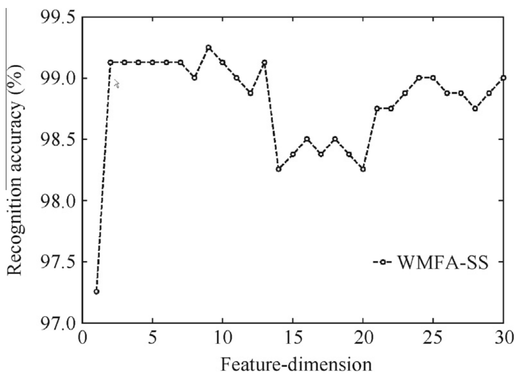


Fig. 5 Recognition accuracy versus feature-dimension on BH-AIR-2.0.

Table 3 Recognition accuracies of WMFA-SS versus MCF-MI.

Approach	Accuracy (%)	Dimension
MCF-MI	92.18	
WMFA-SS	99.25	9

built the BH-AIR-2.0, which was described in the same way with the database used by MCF-MI. As mentioned, the recognition accuracy of MCF-MI is 92.81%, which the aircraft is not covered and 83.35%, which is covered. We implement our proposed algorithm on BH-AIR-2.0, which includes the situation of being covered (Fig. 2).

According to the last experiment, we firstly split the BH-AIR-2.0 with 300 training samples per pattern of aircraft and the rest of them as test samples and then set the optimized parameters of intrinsic graph and penalty graph – k_1 as 30 and k_2 as 400. In penalty graph, the parameter σ of weight function is still $(1/2N_{k_2})\sum||\boldsymbol{x}_i - \boldsymbol{x}_j||^2$, where $(\boldsymbol{x}_i - \boldsymbol{x}_j)$ represents the nearest pairs among the k_2-nearest among the set $\{(i, j), i \in \pi_c, j \notin \pi_c\}$. The parameter of α is set as 0.1.

We show the dimensionality of invariant features from 1 to 30 in Fig. 5 and we choose the dimension of subspace corresponding to the highest recognition accuracy, which is 9 dimension. The accuracy of our proposed algorithm in BH-AIR-2.0 performances better than MCF-MI, which is shown in Table 3.

5. Conclusions

(1) A new algorithm called Weighted Marginal Fisher Analysis with Spatially Smooth (WMFA-SS) for extracting invariant features in aircraft recognition is proposed in this paper. By designing penalty graph with Heat Kernel function to characterize the interclass separability, the problem of classification caused by aircraft's poses variation has been overcome due to better characterizing the interclass relationship.
(2) Using Laplacian penalty to constrain the coefficients of intrinsic graph to be spatially smooth describes the image prior information that neighboring pixels are correlated. Moreover, the regularization effectively avoids the singularity of Laplacian matrix of intrinsic graph and over-fitting for lacking enough training samples.
(3) An open problem in WMFA-SS is the model selection of parameters, which include k_1, k_2, σ, α. The current efficient method to handle this is cross validation. Additional theoretical analysis is needed for this topic in future.

Acknowledgments

The authors are grateful to Prof. Zhang Guangjun for discussions and Ge Guangyi for preparing data. We also thank the anonymous reviewers for their critical and constructive review of the manuscript. This study was co-supported by the National Key Scientific Instrument and Equipment Development Project (No. 2012YQ140032).

References

1. Li W, Bebis G, Bourbakis NG. 3D object recognition using 2D views. *IEEE Trans Image Process* 2008;**17**(11):2236–54.
2. Flusser J, Suk T. Pattern recognition by affine moment invariants. *Pattern Recogn* 1993;**26**(1):167–74.
3. Khotamzad A, Hong YYH. Invariant image recognition by Zernike moments. *IEEE Trans Pattern Anal Mach Intell* 1990;**12**(5):489–97.
4. Xiong Y, Ma HM. Extraction and application of 3D object SIFT feature. *J Image Graphics* 2010;**15**(5):814–9.
5. Zhu XF, Ma CW. Aircraft recognition based on multiple classifier fusion with multiple invariants. *Chin J Sci Instrum* 2011;**32**(7):1620–8 [Chinese].

6. Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Adv Neural Inf Process Syst* 2001;585–91.
7. Roweis ST, Saul LK. Nonlinear dimensional reduction by locally linear embedding. *Science* 2000;**290**(5500):2323–6.
8. Tenenbaum JB, De SV, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science* 2000;**290**(5500):2319–23.
9. Niyogi X. Locality preserving projections. *Adv Neural Inf Process Syst* 2004;**16**:153.
10. He XF, Cai D, Yan SC, Zhang HJ. Neighborhood preserving embedding. In: *Proceedings of IEEE international conference on computer vision*; 2005. p. 1208–13.
11. Yan SC, Xu D, Zhang BY, Zhang HJ. Graph embedding and extensions: a general framework for dimensionality reduction. *IEEE Trans Pattern Anal Mach Intell* 2007;**29**(1):40–51.
12. Yan SC, Xu D, Zhang BY, Zhang HJ. Graph embedding: a general framework for dimensionality reduction. In: *IEEE international conference on computer vision and pattern recognition*; 2005. p. 830–7.
13. Ng AY, Jordan MI, Weiss Y. On spectral clustering: analysis and an algorithm. *Adv Neural Inf Process Syst* 2002;**2**:849–56.
14. Cai D, He XF, Hu YX, Han JW, Thomas H. Learning a spatially smooth subspace for face recognition. In: *IEEE international conference on computer vision and pattern recognition*; 2007. p. 1–7.
15. Martinez AM, Kak AC. PCA versus lda. *IEEE Trans Pattern Anal Mach Intell* 2001;**23**(2):228–33.
16. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. New York: Springer-Verlag; 2001, p. 5.
17. Buzbee BL, Golub GH, Nielson CW. On direct methods for solving poisson's equations. *SIAM J Numer Anal* 1970;**7**(4):627–56.
18. Jost J. *Riemannian geometry and geometric analysis*. Berlin: Springer; 1995.
19. Finbarr OS. Discretized laplacian smoothing by Fourier methods. *J Am Stat Assoc* 1991;**86**(415):634–42.

Wei zhenzhong received the Ph.D. degree in Precise Instrument and mechanism from Beihang University and then became a Professor there. His main research interests are Visual Measurement, Image Processing and Pattern Recognition.

Liu Chang is a postgraduate at school of Instrument and Optical-Electronics, Beihang University. He received his B.S. degree from Beihang University in 2010. His area of research includes Visual Measurement, Image Processing and Machine Learning.